

Applicant(s): JACOBSEN et al.
Appl. No. 10/082,354

Remarks

Applicant thanks the Examiner for the Office Action dated February 22, 2007 (Notice of Non-Compliant Amendment under 37 CFR 1.121). In response, all claim dependency from canceled claim 2 has been amended to dependency from independent claim 1. In addition, other formalities have been corrected as well, namely by addition of the word "storage" in claim 48, and the amendment of punctuation and conjunction language in claims 60 and 61.

Applicant respectfully resubmits the following remarks, which were originally submitted in Applicant's Amendment filed November 22, 2006.

Applicant has carefully reviewed the contents of the Office Action dated August 23, 2006. Reconsideration and allowance of the subject application in view of the foregoing amendments and the following remarks is respectfully requested.

By this Amendment, claim 2 is canceled, claims 54-59 remain withdrawn, and claims 60, 61 are newly added. Accordingly, upon entry of the Amendment, claims 1, 3-53, 60 and 61 are left pending.

Claims 1, 3, 6, 21, 25, 35, 50 and 53 were objected to due to certain informal language. Though Applicant respectfully disagrees with numerous Examiner's assertions, such language has been amended herewith to remove the basis of such objection and to promote the prosecution on the merits.

Claims 1, 14-23, 27, 33-37, 40-43 and 46-48 were rejected under 35 USC § 102(e) as being anticipated by U.S. Patent No. 6,584,468 to Gabriel *et al.* (hereinafter "Gabriel").

Applicant respectfully traverses this rejection in view of the claim amendments and the following remarks.

In Gabriel, the disclosure begins with an example set of documents (or URLs) chosen by a human editor. For example, the Abstract states “[...] the corpus is searched using the set of reference documents to find one or more of the documents in the corpus that contain information in the domain relevant to the query.” Gabriel, in this regard, appears similar to The invention is in this respect very similar to a “focused crawler.”

Gabriel discloses a pattern recognition method and system, meaning it takes aims to collect documents similar to the example document(s) provided to it manually. Refer to col. 3, ll. 47-61. Web pages corresponding to relevant links are parsed and indexed themselves. The process continues until there are no more relevant links or the process is terminated. Accordingly, the Gabriel's method is self-learning, because "the initial seeds used in the master search list are augmented by relevant child links found within the seed web pages." (emphasis added).

Nowhere does Gabriel disclose, hint or suggest any terminology resembling Applicant's recitation of a filter. Rather, Gabriel instead relies on random/incidental terms present in material and queries instead of using a filter. Referring to col. 8, 1.2-4, the content of the HTML file is searched for occurrences of terms, words or phrases “based on a table of relevant or sought after terms.” Accordingly, it is clear that not only does Gabriel not disclose or suggest use of a filter, but rather, teaches away from such use, as he provides the above noted approach.

For example, the method may attempt to determine a pattern from the original samples supplemented with query terms and index these, e.g., a horse is a four legged animal with hoofs, a fairly long neck, a mane and a tail. The method will then attempt to find descriptions of creatures with the same characteristics and will come up with, for example, zebras, giraffes, because the system recognizes neither the concept of a horse nor the context thereof, but merely attempts to guess or approximate, based on the basis of a very fuzzy description deducted from the sample documents. For example, the terms may be more or less suitable, and the quality will most likely be impaired by the query terms that supplements them.

The system accordingly attempts to understand the context and all material stored, qualified to be within the context according to the specific and unique (for example, human) terminology of the context.

Therefore, Gabriel's method and system has little if any correlation to subject specific searches, as claimed by Applicant. Rather it is a pattern recognition method unaware of the subject (or topic) or the context.

Also, nowhere does Gabriel disclose, hint or suggest, as does applicant, employing a terminology/lexicon based type of filtering. As there is no filtering, and as the aforementioned type of filtering applies to terminology/lexicon type of filtering, such filtering could apply to Applicant's claimed method and system, not that disclosed by Gabriel.

Furthermore, Gabriel does not hint, suggest or disclose "filtering the contents of a site at least a second time for relevancy to said particular subject." Nor would the latter be obvious in

view of the Chen reference, for the reasons mentioned below.

As the recitation of claim 2, now canceled, has been incorporated in claim 1, Applicant respectfully submits that recitation of "filtering the contents of a site at least a second time for relevancy to said particular subject" is not hinted at or suggested by Gabriel in view of Chen. The basis for any correlation between Applicant's recitation of filtering and the methods and systems disclosed by either Gabriel or Chen, are clearly refuted based on the following remarks.

Claims 3-13, 24, 25, 45 and 50-53 were rejected under 35 USC § 103(a) as being unpatentable over Gabriel in view of U.S. Patent No. 6,349,307 to Chen (hereinafter "Chen"). Applicant respectfully traverses this rejection in view of the claim amendments and the following remarks.

As noted, Gabriel method and system has little if any correlation to subject specific searches, as claimed by Applicant. Chen fails to overcome this deficiency, as Chen has nothing to do with subject specific searches, but rather is a pattern recognition method unaware of the subject (or topic) or the context.. Chen attempts to optimize the collection of a subset of a document collection, for example, the Web, with respect to an example set of URLs.

In Chen, the description begins with an example set of documents (or URLs) chosen by a human editor. In interpreting and analyzing these documents, the system attempts to extract a pattern of similarities. Based on this pattern, Chen's system will test other documents to see if they match. If a document matches the pattern it will be considered relevant to the topic, whereas if it does not, it will be excluded. Again, the description relates to a focused crawler.

Accordingly, Chen provides a pattern recognition variant, meaning the system aims to collect documents similar to the example documents it has been provided manually. See c.5, ll.5-51. In a training mode, the topic extractor of each node is provided feedback regarding its accuracy in determining each of the topics. The feedback is then used to adjust the topic extractor's subsequent determinations. For example, a Bayesian network is used which adjusts a likelihood factor associated with each decision component, based on the accuracy of each such determination. Whereas correct determinations increase the likelihood factors associated with the decision components, incorrect determinations decrease such factors. Machine learning techniques are similarly applied to create or modify the hierarchical classifications of topics based, for example, on observed clusterings of document topics.

Therefore, it is clear, that Chen's approach *employs no concept of a filter, as presently claimed*, but rather relies on random/incidental terms present in the initial material. Nor does Chen hint or suggest this recitation missing from Gabriel's teaching. From the above, Chen's method will attempt to determine a pattern from the original samples supplemented with query terms etc., and index these. For example, a horse may be characterized as a four legged animal with hoofs, a fairly long neck, a mane and a tail. Chen's method will then attempt to find descriptions of creatures with the same characteristics and will come up with, for example, zebras, giraffes, and the like, because the system understands neither the concept of a horse nor the context, but merely attempts to guess based on the basis of a very fuzzy description deducted from the sample documents. Accordingly, the understanding of the context and all material

Applicant(s): JACOBSEN et al.
Appl. No. 10/082,354

stored qualifies to be within the context according to the specific and unique (human) terminology of the context, but rather is a pattern recognition method unaware of the subject (or topic) or the context.

Accordingly, Chen's method and system fail to compensate for Gabriel's failures, as it too provides a method and system with no correlation to subject specific searches, unlike as claimed by Applicant. Rather it is a pattern recognition method unaware of the subject (or topic) or the context. It also would not have been obvious to persons skilled in the art to shift from a pattern recognition approach to a semantical approach, in view of the foregoing remarks.

Claims 26, 28-32, 38 and 39 were rejected under 35 USC § 103(a) as being unpatentable over Gabriel in view of U.S. Patent No. 6,636,848 to Aridor *et al.* (hereinafter "Aridor"). Applicant respectfully traverses this rejection in view of the claim amendments and the following remarks.

Aridor combines meta-crawling with pattern recognition, like in the foregoing Chen and Gabriel references. Aridor's systems uses "knowledge agents" which are small scripts, i.e., client applications, that start out with an example set of documents (or URLs) chosen by a human editor, as in Chen and Gabriel.

When the user desires to make a search, the user will choose an appropriate agent, such as an agent optimized for the particular topic, by comparing all results with documents like the one the user desires to find. The user will enter the query terms and the agent will take the query to one or more general purpose search engines. The agent will then take the documents

referenced in the results lists and compare them to the sample document sorting out the ones that best match the particular pattern, and return these or references to these to the user by mail or otherwise.

Aridor's system is primarily targeted at "pervasive devices" like PDA's that are not online for a long time at a time. This means that it will attempt to determine a pattern from the original samples supplemented with query terms and index these. The foregoing example in relation to a horse applies as well, where the concept of a horse or the context are not understood, but rather the system merely attempts to guess, based on the basis of a very fuzzy description deducted from the sample documents.

Aridor's system, like Gabriel and Chen, fails to hint or suggest a subject specific search, but rather is a pattern recognition method unaware of the subject (or topic) or the context. It discloses, hints or suggest nothing equivalent to an SSSE system in terms of search technology, where it merely query other search engines but has no crawler, nor does it teach anything about subject specific searches, but only at best searches based on examples, nor does it teach anything about ranking, where it is based on a very crude mechanism of counting words and links respectively.

Accordingly, Chen's method and system fail to compensate for Gabriel's and Chen's failures, as it too provides a method and system with no correlation to subject specific searches, unlike as claimed by Applicant, and further fails to provide Applicant's recited filtering, for the foregoing reasons.

Applicant(s): JACOBSEN et al.
Appl. No. 10/082,354

The Examiner objected to claim 44 as being dependent upon a rejected base claim, but would be allowable if rewritten in independent form, including all of the limitations of the base claim and intervening claims. New claims 60 and 61 respectively represent independent versions of claim 44 in method and system claim formats. Accordingly, Applicant respectfully submits the foregoing claims 60 and 61 are allowable.

Conclusion

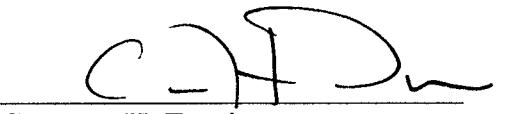
Based on Applicant's above noted arguments, Applicant respectfully submits that the foregoing rejections should be withdrawn. All objections and rejections having been addressed, it is respectfully submitted that the present application is in condition for allowance, and notice to such effect is earnestly solicited. If the Examiner believes, for any reason, that personal communication will expedite prosecution of this application, the Examiner is hereby invited to telephone the undersigned at the number provided.

Please charge any shortage in fees due in connection with the filing of this paper, including extension of time fees, to Deposit Account No. 22-0261.

Applicant(s): JACOBSEN et al.
Appl. No. 10/082,354

Prompt and favorable consideration of this Amendment and Reply is respectfully
requested.

Respectfully submitted,



Date: March 22, 2007

Cameron H. Tousi
Registration No. 43,197
VENABLE LLP
P.O. Box 34385
Washington, D.C. 20043-9998
Telephone: (202) 344-4000
Direct Dial: (703) 760-1913
Telefax: (202) 344-8300